AD_____

Award Number: W81XWH-08-1-0673

TITLE:

Engineer Novel Anticancer Bioagents

PRINCIPAL INVESTIGATOR:

Yi-Qiang Cheng, Ph.D. (Principal Investigator)
Cheng Wang, Ph.D. (Postdoctoral Research Associate)
Shane Wesener, M.S. (Graduate Research Assistant)
Vishwakanth Potharla, M.S. (Graduate Teaching Associate)

CONTRACTING ORGANIZATION:

University of Wisconsin System
Milwaukee, Wisconsin 53211

REPORT DATE:

October 2010

TYPE OF REPORT:

Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

√ Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-10-2010	Annual	15 Sept 2009 – 14 Sept 2010

4. TITLE AND SUBTITLE

Engineer Novel Anticancer Bioagents

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-08-1-0673

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Yi-Qiang Cheng, Ph.D. (Principal Investigator)
Cheng Wang (Postdoctoral Research Associate)
Shane Wesener (Graduate Research Assistant)
Vishwakanth Potharla (Graduate Teaching Assistant – limited support)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Wisconsin System
Milwaukee, Wisconsin 53211

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Material Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During the second contract-year of our research we have successfully reconstituted the entire FK228 biosynthetic gene cluster (~35 kb) plus additional DNA elements (total ~6 kb) into a single integrative vector, a miniHimar RB1 transposon, through complicated manipulations. However, repeated attempts to integrate the gene cluster-carrying transposon into *E. coli* chromosome for stable functioning without antibiotic selection (hence to create marker-free genetically modified organism – GMO as required by FDA regulations) have failed. The overall transformation efficiency is extremely low ($<10^{-9}$ transformants per ug DNA) and always resulted in partial deletion or rearrangement of the gene cluster. We reason that the large size of the gene cluster (total 38 kb, which is still one of the small gene clusters known to produce an FDA-approved anticancer drug) and two internal homologous regions might have contributed to the problem, which was not forseen in the first place. To overcome those problems, we have designed and initiated a completely new strategy which employs a much small gene cluster (~9 kb) that contains only two genes, for the production of a potent anticancer agent tryprostatin B. We are therefore behind the projected schedule of work and we modified the project design. We are confident that we can achieve the project goal but we one year of non-budgetary extension of the entire project duration to end by 09/14/2012.

15. SUBJECT TERMS
FK228 biosynthesis, biosynthetic gene cluster, *E. coli*, reconstitution, heterologous production; tryprostatin B

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	33	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Introduction

Our awarded research project is to engineer the next generation of novel anticancer bio-agents that will target and eradicate a broad spectrum of solid tumors including breast cancer. Our research design was based on the hypotheses that (1) the next generation of bio-agents could be engineered to continuously produce doses of an anticancer drug (specifically FK228) at the sites of bacterial spore germination and vegetative growth in the hypoxic/necrotic regions of solid tumors, and (2) that the synergistic actions of bacterial consumption of tumor mass and anticancer drug activities could destroy many kinds of solid tumors regardless of cancer genotype, without the need for prodrug injections and could overcome the common side effects and less efficacy of current cancer therapies.

The project has three specific aims, corresponding to three tasks defined in the *State of Work*, with each aim/task to be accomplished in each of the three-contract-year duration. Our research in the first contract-year (2008-2009) had achieved the first two slightly modified aims/tasks: (1) we reconstituted the FK228 biosynthetic gene cluster and other necessary DNA elements onto three compatible plasmids, and (2) upon introduction of those plasmids into *E. coli* cells, the engineered *E. coli* strain was able to express the genes and produced FK228 under both aerobic and anaerobic fermentation conditions.

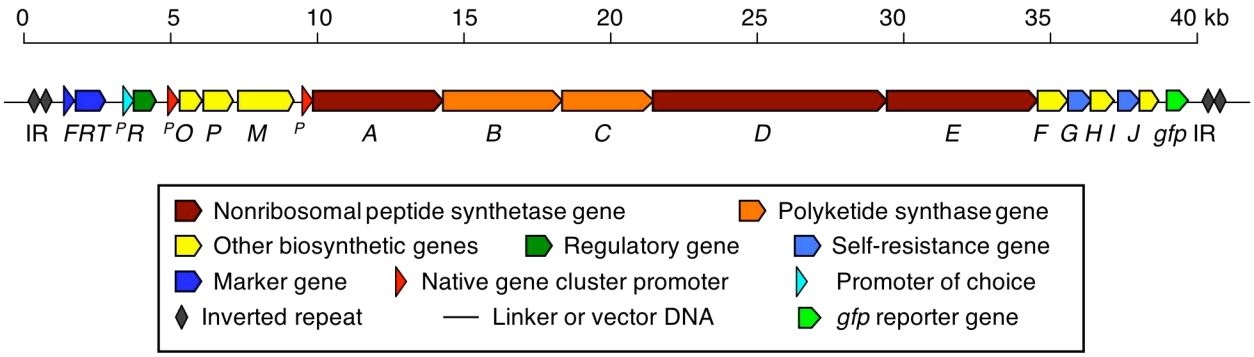
The project goal of the second contract-year (2009-2010) was to integrate the FK228 gene cluster into *E. coli* chromosome for stable functioning without antibiotic selection (hence to create marker-free genetically modified organism – GMO, as required by FDA regulations).

Key Research Status

1. Reconstitution of a complete FK228 biosynthetic gene cluster on a single vector – miniHimar transposon

In order to forcefully integrate a functional FK228 biosynthetic gene cluster into *E. coli* chromosome, first it is necessary to reconstitute the entire biosynthetic gene cluster on a single vector. We chose miniHimar RB1 transposon as the vector for this purpose because it has been reported to be able to integrate randomly into the chromosome of a broad range of bacterial species (Bouhenni, Gehrke et al. 2005; Choi and Kim 2009).

We employed complicated DNA cloning schemes (details omitted for simplicity) and eventually reconstituted a complete FK228 biosynthetic gene cluster from the three plasmids made during the first year of research. This gene cluster includes sequentially a left-end inverted repeat sequence (IR), a removable *FRT* cassette, an interchangeable promoter responsive to anaerobic condition, a stretch of *dep*-genes (*depR*, *depO*, *depP*, *depM*, *depN*, *depA* through *depJ*), a *gfp* reporter gene, and a right-end inverted repeat sequence:



2. Unsuccessful attempts to integrate the engineered FK228 gene cluster into E. coli chromosome

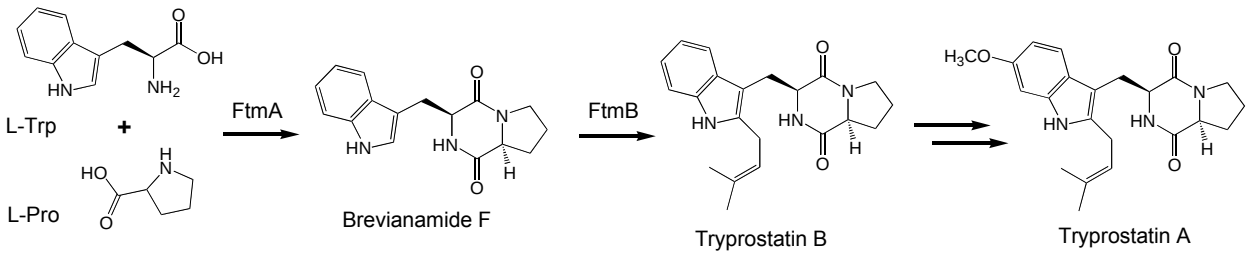
We attempted numerous times to introduce and integrate the FK228 gene cluster-carrying transposon into *E. coli* cells by transformation or electroporation under different conditions (varying amount of DNA, duration of treatment, temperature, antibiotic selection strength, etc.). The transformation efficiency was generally extremely low (<10^{-9} transformants per ug DNA). A few transformants were indeed obtained after extended period of incubation at 37°C. However, upon analysis of the integrated DNA by overlapping PCR amplification, we found that the FK228 gene cluster was either partially deleted or rearranged. Consequentially no FK228 was detected by LC-MS from the fermentation broth of any *E. coli* transformants.

We attributed the failure to the large size of the DNA construct (~45 kb total) and perhaps two known internal homologous regions within the gene cluster.

Facing this insurmountable technical difficulty, we had to decide to suspend this experiment. Meanwhile we searched and conceived an alternative project design, which may circumvent the above technical difficulty (see below).

3. Alternative Project Design: *Engineering and integrating a small tryprostatin B biosynthetic gene cluster (~9 kb) into E. coli chromosome and anaerobic bacterial chromosome*

Tryprostatin A and tryprostatin B are natural products of a fungal species *Aspergillus fumigatus* BM939, with reported bioactivity as potent inhibitors of mammalian cell cycle (Cui, Kakeya et al. 1995; Cui, Kakeya et al. 1996; Cui, Kakeya et al. 1996) and as inhibitors of a breast cancer resistance protein (BCRP) (Woehlecke, Osada et al. 2003; Jain, Zhang et al. 2008). A gene cluster responsible for the biosynthesis of tryprostatins has been characterized (Kato, Suzuki et al. 2009; Maiya, Grundmann et al. 2009) and a model of biosynthesis is proposed as such:



Tryprostatin B is made from L-Trp and L-Pro by two steps of reaction catalyzed by FtmA (encoded by a 6.6-kb gene *ftmA*) and FtmB (encoded by a 1.6-kb gene *ftmB*). Therefore it takes a very small gene cluster (promoter plus *ftmA* and *ftmB*, about 9 kb DNA) to sufficiently produce an anticancer compound tryprostatin B. Two additional enzymes (a cytochrome P450 encoded by ~1.7 kb DNA and O-methyltransferase encoded by ~1.0 kb DNA, respectively) can further convert tryprostatin B to tryprostatin A.

We have obtained the fungal strain, have designed and initiated this alternative project:

(1) To clone both *ftmA* and *ftmB* genes under an *fnr* promoter into a cloning vector pSP72;

(2) To verify the gene fidelity by re-sequencing;

(3) To subclone the entire *ftm* gene cluster (~9 kb) into the miniHimar RB1 transposon;

(4) To attempt to integrate the *ftm* gene cluster into *E. coli* chromosome;

(5) Once we have achieved those milestones, we will exam the productivity of tryprostatin B by the engineering *E. coli* strain. Should the result be positive, we will attempt to introduce and integrate the gene cluster into anaerobic *Bifidobacterium logum* chromosome;

(6) We will then evaluate the antitumor activity of engineered bio-agents.

Reportable Outcome

A manuscript has been favorably reviewed and is under revision for publishing in *Applied and Environmental Microbiology* (see **Appendix**).

Summary

We have achieved partial success toward engineering novel antitumor bio-agents by reconstituting and expressing the FK228 biosynthetic gene cluster in *E. coli* cells. Efforts to integrate the gene cluster into *E. coli* chromosome have suffered serious setback due to insurmountable technical difficulty. For this reason we have designed and initiated an alternative project that will employ a much smaller gene cluster; this gene cluster is responsible for the biosynthesis of tryprostatin B, a potent inhibitor of breast cancer resistance protein (BCRP).

We have confidence to achieve the project goal in about two years from now. Therefore I here request a non-budgetary extension of the entire project duration to end by 09/14/2012.

References

Bouhenni, R., A. Gehrke, et al. (2005). "Identification of genes involved in cytochrome c biogenesis in Shewanella oneidensis, using a modified mariner transposon." Appl Environ Microbiol **71**(8): 4935-4937.

Choi, K. H. and K. J. Kim (2009). "Applications of transposon-based gene delivery system in bacteria." J Microbiol Biotechnol **19**(3): 217-228.

Cui, C. B., H. Kakeya, et al. (1996). "Novel mammalian cell cycle inhibitors, tryprostatins A, B and other diketopiperazines produced by Aspergillus fumigatus. I. Taxonomy, fermentation, isolation and biological properties." J Antibiot (Tokyo) **49**(6): 527-533.

Cui, C. B., H. Kakeya, et al. (1995). "Tryprostatins A and B, novel mammalian cell cycle inhibitors produced by Aspergillus fumigatus." J Antibiot (Tokyo) **48**(11): 1382-1384.

Cui, C. B., H. Kakeya, et al. (1996). "Novel mammalian cell cycle inhibitors, tryprostatins A, B and other diketopiperazines produced by Aspergillus fumigatus. II. Physico-chemical properties and structures." J Antibiot (Tokyo) **49**(6): 534-540.

Jain, H. D., C. Zhang, et al. (2008). "Synthesis and structure-activity relationship studies on tryprostatin A, an inhibitor of breast cancer resistance protein." Bioorg Med Chem **16**(8): 4626-4651.

Kato, N., H. Suzuki, et al. (2009). "Identification of cytochrome P450s required for fumitremorgin biosynthesis in Aspergillus fumigatus." Chembiochem **10**(5): 920-928.

Maiya, S., A. Grundmann, et al. (2009). "Improved tryprostatin B production by heterologous gene expression in Aspergillus nidulans." Fungal Genet Biol **46**(5): 436-440.

Wang, C., S. R. Wesener, et al. (2009). "An FAD-dependent pyridine nucleotide-disulfide oxidoreductase is involved in disulfide bond formation in FK228 anticancer depsipeptide." Chem Biol **16**(6): 585-593.

Woehlecke, H., H. Osada, et al. (2003). "Reversal of breast cancer resistance protein-mediated drug resistance by tryprostatin A." Int J Cancer **107**(5): 721-728.

Appendix

Shane R. Wesener, Vishwakanth Y. Potharla and Yi-Qiang Cheng. Reconstitution of FK228 Biosynthetic Pathway Revealing Cross-Talk between Modular Polyketide Synthases and Fatty Acid Synthase (favorably reviewed and under revision).

Reconstitution of FK228 Biosynthetic Pathway Revealing Cross-Talk between Modular Polyketide Synthases and Fatty Acid Synthase

Shane R. Wesener,[1] Vishwakanth Potharla[1] and Yi-Qiang Cheng[1,2,3]*

[1]Department of Biological Sciences, [2]Department of Chemistry and Biochemistry, University of Wisconsin–Milwaukee, Milwaukee, P.O. Box 413, Wisconsin 53201, USA; [3]School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

* Corresponding author. Department of Biological Sciences
University of Wisconsin–Milwaukee
P.O. Box 413
Milwaukee, WI 53201
Phone: (414) 229-4739
Fax: (414) 229-3926
E-mail: ycheng@uwm.edu

Section: Genetics and Molecular Biology

Running title: FK228 biosynthetic gene cluster reconstitution

Keywords: "AT-less" polyketide synthase, *Chromobacterium violaceum*, fatty acid biosynthesis, FK228, gene cluster reconstitution

ABSTRACT

Functional cross-talk between fatty acid biosynthesis and secondary metabolism has been discovered in several cases in microorganisms, none of them however involves a modular polyketide biosynthetic enzyme. Previously we reported a hybrid modular nonribosomal peptide synthetase-polyketide synthase (PKS) pathway for the biosynthesis of FK228 anticancer depsipeptide in *Chromobacterium violaceum* no. 968. This pathway contains two PKS modules on the DepBC enzymes that lack a functional acyltransferase (AT) domain, and no apparent AT-encoding gene exists within the gene cluster or its vicinity. We report here that, through reconstitution of the FK228 biosynthetic pathway in *E. coli* cells, two essential genes, *fabD1* and *fabD2*, both encoding a putative malonyl CoA acyltransferase component of the fatty acid synthase complex, are positively identified to be involved in FK228 biosynthesis. Either gene product appears sufficient to complement the "AT-less" PKS modules on DepBC for polyketide chain elongation. Concurrently a gene (*sfp*) encoding a putative Sfp-type phosphopantetheinyltransferase was identified to be necessary for FK228 biosynthesis as well. Most interestingly, an engineered *E. coli* strain carrying all necessary genetic components produced significant levels of FK228 under both aerobic and anaerobic cultivation conditions. Discovery of the *trans* complementation of modular PKSs by housekeeping ATs reveals natural product biosynthesis diversity. Moreover, demonstration of anaerobic production of FK228 by an engineered facultative bacterial strain validates our effort toward engineering of novel tumor-targeting bio-agents.

INTRODUCTION

Polyketides and nonribosomal peptides constitute two large families of natural products that are biosynthesized by polyketide synthases (PKSs) or nonribosomal peptide synthetases (NRPSs), respectively (16, 24). There are also increasing number of natural products identified as hybrid polyketide-nonribosomal peptide molecules that are biosynthesized by a combination of PKS and NRPS enzymes (14, 29). PKSs function by carrying out successive rounds of decarboxylative Claisen condensation of short carboxylic acid moieties derived from malonyl coenzyme A (CoA), methylmalonyl CoA and a few other less common acyl CoAs (5, 34). In parallel, NRPSs condense amino acids or amino acid derivatives by forming amide bond between two building blocks (15).

PKSs share many similarities with fatty acid synthases (FASs) and can be largely classified as type I, II and III, according to their architecture and mode of catalysis; yet many variants of PKSs exit in nature (5, 32, 34). Type I PKSs are multifunctional enzymes organized into modules, each of which harbors a set of distinct domains responsible for the catalysis of one cycle of polyketide chain elongation. A prototypical type I PKS elongation module contains minimally three integral domains – a ketosynthase (KS), an acyltransferase (AT) and an acyl carrier protein (ACP) – that together catalyze one round of polyketide chain elongation. Optional domains (such as ketoreductase [KR], dehydratase [DH], enoylreductase [ER]) are found between the AT and ACP domains, which carry out variable steps of reductive modifications on polyketide intermediate. Type I PKSs can be subdivided into iterative and noniterative, depending on whether the same set of domains is used once or multiple times during the biosynthesis of a polyketide molecule. Type II PKSs are multienzyme complexes that carry a single set of catalytic domains acting iteratively. Type III PKSs are chalcone synthase (CHS)-

like enzymes that essentially are iteratively acting condensing enzymes. FASs, and type I and type II PKSs all use ACPs to tether acyl CoA substrates during fatty acid or polyketide biosynthesis, whereas type III PKSs directly condense acyl CoA substrates without carrier proteins. In parallel, NRPSs use peptidyl carrier proteins (PCPs) to tether aminoacyl CoA substrates during nonribosomal peptide biosynthesis. Prior to biosynthesis, apo form of ACPs or PCPs must be activated to holo form by attaching a 4'-phosphopantetheinyl moiety from CoA onto a conserved serine residue; this posttranslational modification reaction is catalyzed by phosphopantetheinyltransferases (PPTases) that can be largely classified into two groups, AcpS-type and Sfp-type (12, 21, 26).

Integral AT domains in the prototypical type I PKSs were termed cognate ATs (8). A distinct variant of the noniterative type I PKSs contains no intact cognate ATs but a short segment of remnant AT sequence in some or all modules. This subclass of type I PKSs was named the "AT-less" type I PKSs and the remnant AT segment the AT docking domain (7, 8, 36). The essential AT activities are provided *in trans* by discrete AT enzymes encoded by genes that are physically separated from the PKS genes, but nevertheless within the same gene cluster.

Two unusual natural product biosynthetic gene clusters responsible for the biosynthesis of actinorhodin (28) or tetracenomycin (4, 17, 35) encode type II PKSs that lack an AT subunit and no AT-encoding gene is present within the gene cluster boundaries. The AT activity necessary for polyketide chain extension is, instead, provided by a malonyl CoA acyltransferase (MCAT, also known as FabD, encoded by the *fabD* gene) of type II fatty acid synthase complex (FAS II). Recruitment of an FAS II component for polyketide biosynthesis establishes a functional cross-talk between the bacterial primary metabolism and secondary metabolism.

FK228 is an interesting bicyclic depsipeptide natural product recently approved by the US

Food and Drug Administration (FDA) as a new class of anticancer drug for the treatment of cutaneous T-cell lymphoma (CTCL) (1). FK228 epigenetically affects tumorigenesis by inhibiting histone deacetylases (18). Our research interests reside on the mechanism by which FK228 is biosynthesized and on metabolic engineering of the FK228 biosynthetic pathway for generating structural analogs for drug discovery. Previously we have reported cloning of the FK228 biosynthetic (*dep*) gene cluster and characterization of a critical gene (*depH*) responsible for disulfide bond formation as the final step in FK228 biosynthesis in *Chromobacterium violaceum* no. 968 (9, 37). In a companion paper we further described a precise determination of the *dep* gene cluster boundaries through systematic gene mutation and transcriptional analysis (27). As a result, the *dep* gene cluster is now redefined to contain 12 genes: *depA* through *depJ*, *depM*, and a newly identified pathway regulatory gene *depR*.

There are still several unanswered questions regarding the FK228 biosynthetic pathway. In particular, the two PKS modules on DepBC enzymes do not contain any cognate AT domain but a remnant AT docking domain (Fig. 1c), and there is no AT-encoding gene anywhere in the defined *dep* gene cluster (9, 27). Where does the AT activity necessary for FK228 biosynthesis come from? In this work we report the identification of two essential genes, *fabD1* and *fabD2*, both encoding a putative MCAT component of the FAS II complex, that are positively involved in FK228 biosynthesis. Concurrently we identified a gene (*sfp*) encoding a putative Sfp-type PPTase necessary for FK228 biosynthesis as well. Furthermore, we show that, through reconstitution of the FK228 biosynthesis pathway, FK228 could be produced by recombinant *E. coli* strains under both aerobic and anaerobic cultivation conditions. Our studies revealed, for the first time, that modular PKSs recruit FabD components of the primary metabolism for the biosynthesis of a secondary metabolite. Discovery of the *trans* complementation of modular

PKSs by housekeeping ATs reveals natural product biosynthesis diversity. Moreover, demonstration of anaerobic production of FK228 by an engineered facultative bacterial strain validates our effort toward engineering of novel tumor-targeting bio-agents (6).

MATERIALS AND METHODS

Bacterial strains, plasmids, culture conditions, and general molecular biological manipulations. The bacterial strains and plasmids used in this study are listed in Table 1. Culture conditions and general molecular biological manipulations were performed as described (9, 37), or according to standard protocols (30).

Rapid genome sequencing and gene identification. Genomic DNA of the wild type *C. violaceum* no. 968 strain was prepared from an overnight culture with an UltraClean Microbial DNA Isolation kit (MO BIO Labs, Carlsbad, CA), and was submitted for shotgun single-end and paired-end shotgun sequencing on a G20 FLX platform (454 Life Science, Branford, CT) at the Research Technology Support Facility of Michigan State University (East Lansing, MI). Standard *de novo* assembly of sequence reads was performed by the facility personnel and a draft genome sequence was provided to us as the end users. Candidate genes were identified by using known protein sequences as bait to search the draft genome sequence of *C. violaceum* by Blastx algorithm (3).

General strategies for targeted gene deletion in *C. violaceum* no. 968. A multiplex PCR method, as described elsewhere (10, 27, 37), was used for all intended gene deletion experiments. This method utilized a broad host-range Flp-*FRT* recombination system for site-specific gene replacement/deletion and subsequent marker removal (10, 19). Primers used for making gene deletion constructs and for detection of genotypes are listed in Table S1.

Reconstitution of FK228 biosynthetic gene cluster in engineered *E. coli* strains. To probe whether individual candidate AT-encoding genes or PPTase-encoding genes are involved in FK228 biosynthesis, a three-plasmid system was utilized for gene expression and FK228 biosynthesis in bacterial strains derived from *E. coli* BL21(DE3) (Fig. 2a). First, all candidate genes were amplified by high fidelity PCR from the genomic DNA of *C. violaceum* with primer sets carrying designed restriction sites. Second, two candidate PPTase-encoding gene amplicons were double-digested with BamHI/HindIII and individually cloned into the first multiple cloning site (MCS) of pCDFDuet-1 to create two intermediate constructs, pCDFDuet-1-*acpS* and pCDFDuet-1-*sfp*. Third, two candidate AT-encoding gene amplicons were double-digested with NdeI/KpnI and individually cloned into the second multiple cloning site of pCDFDuet-1 to generate pCDFDuet-1-*fabD1* and pCDFDuet-1-*fabD2*, or into the two previously made intermediate constructs to create a combination of 4 final constructs, pCDFDuet-1-*acpS-fabD1*, pCDFDuet-1-*acpS-fabD2*, pCDFDuet-1-*sfp-fabD1*, pCDFDuet-1-*sfp-fabD2* (Table 1). Finally, the Cosmid 18, which carries the original incomplete FK228 biosynthetic gene cluster that lacks any AT-encoding gene or PPTase-encoding gene (Cheng et al, 2007), and pBMTL-3-*depR*, which was created to complement a *depR*-deletion mutant of *C. violaceum* (27), were used in combination with the above expression constructs for the transformation of *E. coli* BL21(DE3) cells to create a series of bacterial strains (Table 1). Kanamycin at 25 μg/ml, chloramphenicol at 25 μg/ml, and streptomycin at 25 μg/ml were used individually or in combination for selection and maintenance of respective *E. coli* strains.

Bacterial fermentation and quantification of FK228 production by liquid chromatography-mass spectrometry (LC-MS). Wild type *C. violaceum* strain and recombinant *E. coli* strains were fermented aerobically for 4 days at 30ºC under constant

agitation (200 rpm) in 50 ml of LB media supplemented with 1% (w/v) Diaion HP-20 resin (Sigma-Aldrich, St. Louis, MO) and appropriate antibiotics where necessary. Gene expression was induced with 0.5% (w/v) lactose and 0.1 mM IPTG when bacterial culture reached an OD_{600} of 0.4. Strict anaerobic fermentation of bacterial strains was carried out similarly for 5 days at room temperature in a Coy anaerobic chamber (Grass Lake, MI) with occasional manual agitation, except for that 0.05% (w/v) thioglycolate was added to the media to capture any oxidative species generated during fermentation. Extraction of metabolites, and detection and quantification of FK228 by LC-MS were performed as described (27).

RNA extraction and reverse transcription (RT)-PCR. Recombinant *E. coli* BL21(DE3) strains were grown in LB media supplemented with appropriate antibiotics at 30ºC under constant agitation (200 rpm) to an OD_{600} of 0.4. Five ml of each pre-induction sample was collected and the remaining cultures were induced with 0.5% (v/v) lactose and 0.1 mM IPTG. Aliquots of sample were collected at 60 min and again at 120 min post induction. Preservation of sample aliquots, extraction of total RNA, and RT-PCR experiments were performed as described (27). Primers used for detection of individual gene expression are listed in Table S1.

Nucleotide sequence accession numbers. The nucleotide sequences of *C. violaceum* no. 968 genes reported in this paper have been deposited in the GenBank database under accession numbers HM449690 for 16S rRNA gene, HM449691 for *fabD1*, HM449692 for *fabD2*, HM449693 for *AT3*, HM449694 for *acpS*, and HM449695 for *sfp*, respectively.

RESULTS

Draft genome sequencing of *Chromobacterium violaceum* no. 968. Shotgun sequencing of *C. violaceum* genomic DNA on a GS20 FLX Sequencer generated a total of 163,954,650 input

bases, which were assembled into 122 contigs; among them 82 are large contigs (>500 nt) with an average contig size of 59,514 bps. Those contigs were further aligned into 15 scaffolds with a total length of 4,909,141 bps. The sequence coverage for this draft bacterial genome was thus calculated at 33.4 folds, which exceeded the desired 30-fold oversampling of raw sequence for the 454 pyrosequencing and de novo assembly technology platform (13). Compared to the published 4.75-Mb complete genome of a type strain of *C. violaceum* ATCC 12472 with a (G + C) content of 64.8% (11), the 4.91-Mb draft genome sequence of *C. violaceum* no. 968 with an overall (G + C) content of 61.9% obtained in this study appears to be near complete. The quality of this draft genome sequence was assessed by the following analyses. First, a comparison of the 16S rDNA sequences of two *C. violaceum* strains revealed an overall 96% identity without a single gap (Fig. S1), indicating a high quality of the draft genome sequence and a taxonomical relatedness of the two strains. Second, a homology search by Blastn algorithm of the draft genome sequence using our previously published FK228 biosynthetic gene cluster sequence (GenBank no. EF210776) (9) as bait identified two contigs that carry the gene cluster with a 100% sequence identify and with a 1298-bp sequence gap (data not shown). This gap was artificially created due to two highly homologous regions within the gene cluster that were assembled into one copy of sequence.

Identification and initial characterization of candidate genes. Sequences of three bait proteins, FabD of *E. coli* K-12 (GenBank no. AAC74176), FabD of *C. violaceum* ATCC 12472 (GenBank no. NP_903085) and LnmG of *Streptomyces atroolivaceus* S-140 (GenBank no. AAN85520), were used to search the draft genome sequence and identified three candidate genes, *fabD1*, *fabD2* and *AT3*, that encode putative AT enzymes (FabD, MCAT or AT) (Table 2; Fig. 1b). Based on bioinformatics analysis, the *fabD1* gene lies within an apparent FAS II gene

cluster that also includes *fabH*, *fabG*, *acp* and *fabF*. The *fabD2* gene is a standalone gene whose function cannot be predicted *a priori*. The *AT3* gene lies within a putative gene cluster that may involve in cell-surface O-antigen biosynthesis.

Similarly, sequences of five bait proteins, AcpS of *E. coli* K-12 (GenBank no. P24224), Sfp of *Bacillus subtilis* (GenBank no. P39135), AcpS of *C. violaceum* ATCC 12472 (GenBank no. NP_901742), EntD of *C. violaceum* ATCC 12472 (GenBank no. NP_902320) and PcpS of *Pseudomonas aeruginosa* PAO1 (GenBank no. AAG04554), were used to search the draft genome sequence and identified two candidate genes, *acpS* and *sfp*, that encode putative PPTase enzymes (AcpS; Sfp) (Table 2; Fig. 1a). The *acpS* gene appears to be a standalone gene which is likely involved in primary metabolite (e.g. fatty acid) biosynthesis. The *sfp* gene is located at the end of an apparent NRPS gene cluster, therefore it is likely involved in secondary metabolite (e.g. nonribosomal peptides or hybrid molecules) biosynthesis.

Individual candidate genes were subjected to targeted gene deletion by an established multiplex PCR procedure to probe whether they play any roles in FK228 biosynthesis. Gene *AT3* was successfully mutated and the mutant strain did not show any notable difference from the wild type strain (data not shown), suggesting that *AT3* is disposable and is independent of FK228 biosynthesis. Thus *AT3* was not further tested. In contrast, four other candidate genes, *fabD1*, *fabD2*, *acpS* and *sfp*, could not be mutated despite numerous attempts, indicating that they are all essential to the bacterial physiology and survival. They were then subjected to tests by a different strategy described below.

Reconstitution of the FK228 biosynthetic pathway in *E. coli*. A series of recombinant *E. coli* strains were created herein (Table 1; Fig. 2a) and the relative levels of FK228 production by these strains were examined by LC-MS (Fig. 3; Table S2; Fig. S2). When Cosmid 18, the

originally identified large construct that carries the *dep* gene cluster which definitely lacks an AT-encoding gene and a PPTase-encoding gene (9), was introduced into *E. coli* BL21(DE3) cells, the recombinant strain SW01 failed to produce FK228. When a newly defined pathway regulatory gene, *depR* (27), was introduced into SW01 via a broad-host range construct pBMTL-3-*depR*, the resulting strain SW02 still did not produce any detectable level of FK228. When the candidate *sfp* gene was introduced into SW02 via a compatible construct pCDFDuet-1-*sfp*, the resulting strain SW03 produced a moderate level of FK228. When either the candidate *fabD1* or *fabD2* gene was added in tandem with *sfp* on pCDFDuet-1-*sfp-fabD1* or pCDFDuet-1-*sfp-fabD2* construct, the resulting strains of SW07 and SW08 produced significant levels of FK228. All other strains that received a single candidate gene of either *acpS* (strain SW04), *fabD1* (strain SW05) or *fabD2* (strain SW06), or two genes with *acpS* in tandem with either *fabD1* (strain SW9) or *fabD2* (strain SW10) did not show production of detectable level of FK228. Those observations led to the following conclusions: (1) the *sfp* gene but not the *acpS* gene is capable of converting all PCPs and ACPs in the FK228 biosynthetic pathway from their inactive apo form to their active holo form; (2) indigenous *E. coli* AcpS cannot promiscuously act on those heterologous carrier proteins from a secondary metabolic pathway; (3) either gene product of the essential *fabD1* or *fabD2* gene of *C. violaceum* no. 968 is able to complement the "AT-less" PKS modules on DepBC for polyketide chain extension; (4) even *E. coli* FabD (in the background in case of strain SW03) is able to provide the necessary AT activities for FK228 biosynthesis albeit at a lower rate; (5) at this point we believe that genes necessary for FK228 biosynthesis have all been identified.

RT-PCR verification of gene expression. To verify that key biosynthetic and regulatory genes for FK228 biosynthesis were adequately expressed in the SW07 recombinant strain under

normal aerobic conditions, aliquots of bacterial culture were collected at three time points, total RNA samples were prepared and subjected to semi-quantitative RT-PCR analysis (Fig. 2b). Prior to chemical induction, all examined genes including two representative structural genes (*depA* and *depJ*) but excluding the 16S rDNA control were not expressed. This suggests that the *depR* gene in its native position on Cosmid 18 is not functioning in *E. coli* cells, likely due to a lack of proper external or internal stimulus, or signal transduction pathway; expression of the *dep* gene cluster carried by Cosmid 18 in *E. coli* cells requires an ectopic copy of *depR* driven by lactose-inducible promoter on a pBMTL-3 vector. Upon induction by lactose and IPTG for the expression of *depR* from pBMTL-3-*depR* and for the expression of *sfp* and *fabD1* from pCDFDuet-1-*sfp-fabD1*, those genes and two representative structural genes (*depA* and *depJ*) were found to have expressed 60 min post induction; and the gene expression reached higher levels 120 min post induction.

Production of FK228 under anaerobic fermentation conditions. As a pilot study toward engineering of novel tumor-targeting bacterial agents that may effectively infiltrate, multiply and continuously produce an anticancer drug inside the hypoxic core of solid tumors (6), we examined whether the engineered SW07 strain may produce FK228 under anaerobic fermentation conditions. To our delight the SW07 strain produced c.a.14.1% as much FK228 as did by the wild type *C. violaceum* strain or 15.3% by the SW07 strain itself under aerobic conditions (Fig. 3; Table S2; Fig. S2); this relative level of FK228 production was translated into 0.40 mg/L of actual yield or c.a. 741 nM concentration when normalized by FK228 standard. As a control, the wild type *C. violaceum* strain was also found to have produced a lower level of FK228 under anaerobic conditions. This Gram-negative bacterial species appeared to be able to survive and grow slightly during the first 24 hr under the strict anaerobic conditions tested.

DISCUSSION

Organizational and functional variation of PKSs leads to diverse structures of polyketide natural products (5, 32). The modular "AT-less" PKSs represent a severe deviation from the canonical type I PKSs (7, 8) and this phenomenon appears to be a transition state of complex enzyme evolution from FAS II to type II PKSs (33). Nevertheless in most "AT-less" PKS pathways the AT activities required for polyketide chain elongation are often encoded by discrete genes within the respective biosynthetic gene clusters. In the present work, for the first time, we uncovered an extreme case of "AT-less" PKS system where the AT activity is encoded by *fabDs*, the essential bacterial genes involved in fatty acid biosynthesis. Support for this conclusion came from several lines of evidence.

First, sequence analysis of the two PKS modules on DepBC did not find an intact integral AT domain but remnant of AT-docking segment (9); thus DepBC enzymes can be classified as modular "AT-less" PKSs. Further analysis failed to identify any gene within or in the vicinity of the redefined *dep* gene cluster (27). To address the mystery about where the necessary AT activities would come from, the genome of *C. violaceum* no. 968 was decoded by a rapid genome sequencing platform, revealing three candidate genes that may encode the AT activities for FK228 biosynthesis. However, attempts to identify the exact AT-encoding gene by a direct gene deletion approach was not successful, because two candidate genes (*fabD1* and *fabD2*) could not be mutated due to their apparent essential roles in bacterial viability. A third gene (*AT3*) could be mutated but the mutant did not show any obvious sign of physiological defect or decrease in FK228 production.

We then engineered a series of recombinant *E. coli* strains that harbor one to three compatible vectors that carry either the previously identified *dep* gene cluster or candidate genes identified in this work with three objectives in mind (Fig. 2a). First, we hoped to identify the missing AT-encoding gene and PPTase-encoding gene necessary for FK228 biosynthesis. Second, we hoped to reconstitute a functional FK228 biosynthetic pathway in a heterologous host such as *E. coli* cells. Third, we hoped to demonstrate that FK228 could be produced by a recombinant *E. coli* strain under both aerobic conditions and particularly anaerobic conditions as a pilot experiment for genetic engineering of novel cancer-targeting bio-agents (6).

Three engineered *E. coli* strains, SW03, SW07 and SW08, were subsequently found to be able to produce variable amounts of FK228 under normal (aerobic) fermentation conditions (Fig. 3; Table S2; Fig. S2). In addition to the Cosmid 18 and the *depR* gene on a vector, all these three strains received the *sfp* gene; while other *E. coli* strains, regardless of having the *acpS* gene or not, did not produce FK228. It was thus clear that the AcpS-type PPTase from either *E. coli* host or from *C. violaceum* cannot activate the carrier proteins from the FK228 biosynthetic pathway. Therefore it was concluded that the *sfp* gene, which was predicted to encode a broad substrate-range Sfp-type PPTase, is involved in FK228 biosynthesis. This PPTase should have other essential function as well, otherwise the *sfp* gene could have been mutated in the first place. In addition, two of the three strains, SW07 and SW08, which also received either *fabD1* or *fabD2* gene of *C. vioalceum*, produced much higher levels of FK228 than did by SW03. Those observations suggested that either *fabD1* or *fabD2* of *C. violaceum* is involved in FK228 and their involvement is interchangeable; it is also possible that both genes are redundantly involved in FK228 biosynthesis. Surprisingly, the indigenous *fabD* of *E. coli* appeared to function as well for FK228 biosynthesis in the SW03 strain which received the *sfp* gene but not *fabD1* or *fabD2*.

Sequence alignment of the *C. violaceum* FabD1 and FabD2, and *E . coli* FabD shows that FabD1 and FabD share a much higher amino acid sequence identity than other two sets of comparison (Fig. S3), suggesting that FabD1 is very likely the housekeeping FAS II component for bacterial fatty acid biosynthesis. The role of FabD2 is less certain. The *fabD2* gene was positively identified to be involved in FK228 biosynthesis; but like the Sfp-type PPTase, FabD2 should have other essential function as well, otherwise the *fabD2* gene could have been mutated in early experimental attempts.

Now it is proven that an MCAT component of FAS II is recruited by PKS modules on DepBC for FK228 biosynthesis, which establishes a functional cross-talk between the bacterial primary metabolism and secondary metabolism, and adds new evidence to the phenomenon of complex enzyme evolution from FAS II to type II PKSs (33). Besides the cases where an ACP component of FAS II is recruited by a type II PKS system for the biosynthesis of actinorhodin (28) or tetracenomycin (4, 17, 35), other types of cross-talk have also been reported. For example, in the biosynthesis of quinoxaline antibiotics in two *Streptomyces* strains, an ACP from the primary FAS II is recruited by a standalone condensation domain of type II NRPS to form an initiation module of the biosynthetic pathway (31). Furthermore, cross-talk between a type I FAS (FAS I) for primary metabolism and a secondary FAS I for HC-toxin biosynthesis had also be postulated in the fungal species *Cochliobolus carbonum* where only one gene encoding the β-subunit of FAS I was found in the secondary metabolic gene cluster, while the α-subunit of FAS I necessary for HC-toxin biosynthesis might be recruited from the primary FAS I (2). Recently an opposite example was reported in the apicidin biosynthetic pathway in *Fusarium semitectum* where a gene encoding the α-subunit of FAS I is present in the apicidin biosynthetic gene cluster while the β-subunit of FAS I necessary for apicidin biosynthesis is speculated to be recruited

from the primary FAS I (20).

Lastly, the ability by which the engineered SW07 strain produced 0.4 mg/L of actual yield or c.a. 741 nM concentration of FK228 under strict anaerobic conditions is worth highlighting (Fig. 3; Table S2; Fig. S2). Aimed at engineering of tumor-targeting bio-agents (6), this accomplishment represents an important milestone toward the goal. The next step will be undertaken to introduce the FK228 biosynthetic capacity into one or more selected nonpathogenic anaerobic bacterial species, such as *Bifidobacterium longum* or *Clostridium oncolyticum* (25, 38), so that when the engineered bacteria infiltrate and multiply inside the necrotic region of solid tumors, FK228 will be produced *in situ*. Because FK228 is a potent HDAC inhibitor and a new anticancer drug effective in the lower nM range (22, 39), the synergistic efforts of bacterial oncolysis of tumor tissue and the anticancer activity of FK228 could potentially be clinically effective against many types of cancer.

ACKNOWLEDGEMENTS

We thank the Research Technology Support Facility of Michigan State University for providing quality shotgun genome sequencing and assembly services. We also thank Daad Saffarini for assistance on bacterial anaerobic fermentation and Patrick Anderson on LC-MS.

This work was supported in part by a Research Growth Initiative Award from the University of Wisconsin-Milwaukee and by an Idea Award BC073985 from the US Department of Defense Breast Cancer Research Program.

REFERENCES

1. ____. 2010. StatBite: FDA oncology drug product approvals in 2009. J. Natl. Cancer Inst. **102:**219.
2. **Ahn, J. H., and J. D. Walton.** 1997. A fatty acid synthase gene in *Cochliobolus carbonum* required for production of HC-toxin, cyclo(D-prolyl-L-alanyl-D-alanyl-L-2-amino-9, 10-epoxi-8-oxodecanoyl). Mol. Plant Microbe Interact. **10:**207-14.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403-10.
4. **Bao, W., E. Wendt-Pienkowski, and C. R. Hutchinson.** 1998. Reconstitution of the iterative type II polyketide synthase for tetracenomycin F2 biosynthesis. Biochemistry **37:**8132-8.
5. **Chan, Y. A., A. M. Podevels, B. M. Kevany, and M. G. Thomas.** 2009. Biosynthesis of polyketide synthase extender units. Nat. Prod. Rep. **26:**90-114.
6. **Cheng, Y.-Q.** 2008-2011. A specific project supported in part by a University of Wisconsin–Milwaukee Research Growth Initiative Award and by an Idea Award BC073985 from the US Department of Defense Breast Cancer Research Program, University of Wisconsin–Milwaukee, Milwaukee, WI 53201.
7. **Cheng, Y. Q., J. M. Coughlin, S. K. Lim, and B. Shen.** 2009. Type I polyketide synthases that require discrete acyltransferases. Methods Enzymol. **459:**165-86.
8. **Cheng, Y. Q., G. L. Tang, and B. Shen.** 2003. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. Proc. Natl. Acad. Sci. U S A **100:**3149-54.

9. **Cheng, Y. Q., M. Yang, and A. M. Matter.** 2007. Characterization of a gene cluster responsible for the biosynthesis of anticancer agent FK228 in *Chromobacterium violaceum* no. 968. Appl. Environ. Microbiol. **73:**3460-9.
10. **Choi, K. H., and H. P. Schweizer.** 2005. An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. BMC Microbiol. **5:**30.
11. **Consortium, B. N. G. P.** 2003. The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability. Proc. Natl. Acad. Sci. U S A **100:**11660-5.
12. **Copp, J. N., and B. A. Neilan.** 2006. The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. Appl. Environ. Microbiol. **72:**2298-305.
13. **Droege, M., and B. Hill.** 2008. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. J. Biotechnol. **136:**3-10.
14. **Du, L., and B. Shen.** 2001. Biosynthesis of hybrid peptide-polyketide natural products. Curr. Opin. Drug Discov. Devel. **4:**215-28.
15. **Finking, R., and M. A. Marahiel.** 2004. Biosynthesis of nonribosomal peptides1. Annu. Rev. Microbiol. **58:**453-88.
16. **Fischbach, M. A., and C. T. Walsh.** 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem. Rev. **106:**3468-96.

17. **Florova, G., G. Kazanina, and K. A. Reynolds.** 2002. Enzymes involved in fatty acid and polyketide biosynthesis in *Streptomyces glaucescens*: role of FabH and FabD and their acyl carrier protein specificity. Biochemistry **41:**10462-71.

18. **Furumai, R., A. Matsuyama, N. Kobashi, K. H. Lee, M. Nishiyama, H. Nakajima, A. Tanaka, Y. Komatsu, N. Nishino, M. Yoshida, and S. Horinouchi.** 2002. FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases. Cancer Res. **62:**4916-21.

19. **Hoang, T. T., R. R. Karkhoff-Schweizer, A. J. Kutchma, and H. P. Schweizer.** 1998. A broad-host-range Flp-*FRT* recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene **212:**77-86.

20. **Jin, J. M., S. Lee, J. Lee, S. R. Baek, J. C. Kim, S. H. Yun, S. Y. Park, S. Kang, and Y. W. Lee.** Functional characterization and manipulation of the apicidin biosynthetic pathway in *Fusarium semitectum*. Mol. Microbiol. **76:**456-466.

21. **Lambalot, R. H., A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, and C. T. Walsh.** 1996. A new enzyme superfamily - the phosphopantetheinyl transferases. Chem. Biol. **3:**923-36.

22. **Lane, A. A., and B. A. Chabner.** 2009. Histone deacetylase inhibitors in cancer therapy. J. Clin. Oncol. **27:**5459-68.

23. **Lynch, M. D., and R. T. Gill.** 2006. Broad host range vectors for stable genomic library construction. Biotechnol. Bioeng. **94:**151-8.

24. **Meier, J. L., and M. D. Burkart.** 2009. The chemical biology of modular biosynthetic enzymes. Chem. Soc. Rev. **38:**2012-45.

25. **Minton, N. P.** 2003. Clostridia in cancer therapy. Nat. Rev. Microbiol. **1:**237-42.

26. **Mootz, H. D., R. Finking, and M. A. Marahiel.** 2001. 4'-phosphopantetheine transfer in primary and secondary metabolism of *Bacillus subtilis*. J. Biol. Chem. **276:**37289-98.

27. **Potharla, V., S. R. Wesener, and Y.-Q. Cheng.** 2010. Systematic gene mutation and transcriptional analysis revealing unexpected genetic organization and pathway regulation of the FK228 biosynthetic gene cluster in *Chromobacterium violaceum* no. 968. (submitted).

28. **Revill, W. P., M. J. Bibb, and D. A. Hopwood.** 1995. Purification of a malonyltransferase from *Streptomyces coelicolor* A3(2) and analysis of its genetic determinant. J. Bacteriol. **177:**3946-52.

29. **Richter, C. D., D. Nietlispach, R. W. Broadhurst, and K. J. Weissman.** 2008. Multienzyme docking in hybrid megasynthetases. Nat. Chem. Biol. **4:**75-81.

30. **Sambrook, J., and D. W. Russell.** 2000. Molecular Cloning: a laboratory manual, Third edition ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

31. **Schmoock, G., F. Pfennig, J. Jewiarz, W. Schlumbohm, W. Laubinger, F. Schauwecker, and U. Keller.** 2005. Functional cross-talk between fatty acid synthesis and nonribosomal peptide synthesis in quinoxaline antibiotic-producing streptomycetes. J. Biol. Chem. **280:**4339-49.

32. **Shen, B.** 2003. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. Curr. Opin. Chem. Biol. **7:**285-95.

33. **Smith, J. L., and D. H. Sherman.** 2008. Biochemistry. An enzyme assembly line. Science **321:**1304-5.

34. **Staunton, J., and K. J. Weissman.** 2001. Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. **18:**380-416.

35. **Summers, R. G., A. Ali, B. Shen, W. A. Wessel, and C. R. Hutchinson.** 1995. Malonyl-coenzyme A:acyl carrier protein acyltransferase of *Streptomyces glaucescens*: a possible link between fatty acid and polyketide biosynthesis. Biochemistry **34:**9389-402.

36. **Tang, G. L., Y. Q. Cheng, and B. Shen.** 2004. Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. Chem. Biol. **11:**33-45.

37. **Wang, C., S. R. Wesener, H. Zhang, and Y. Q. Cheng.** 2009. An FAD-dependent pyridine nucleotide-disulfide oxidoreductase is involved in disulfide bond formation in FK228 anticancer depsipeptide. Chem. Biol. **16:**585-93.

38. **Wei, M. Q., K. A. Ellem, P. Dunn, M. J. West, C. X. Bai, and B. Vogelstein.** 2007. Facultative or obligate anaerobic bacteria have the potential for multimodality therapy of solid tumours. Eur. J. Cancer **43:**490-6.

39. **Yoo, C. B., and P. A. Jones.** 2006. Epigenetic therapy of cancer: past, present and future. Nat. Rev. Drug Discov. **5:**37-50.

FIGURE LEGENDS

FIG 1. Identification of missing genes necessary for FK228 biosynthesis in *Chromobacterium violaceum* no. 968. (**a**) Local genetic map of two candidate PPTase-encoding genes, with the *sfp* gene product postulated to participate in phosphopantetheinylation of carrier proteins of the FK228 biosynthetic pathway. (**b**) Local genetic map of three candidate AT-encoding genes, with the *fabD1 and fabD2* gene products postulated to provide AT activities to complement the "AT-less" PKS modules on DepBC proteins for FK228 biosynthesis. (**c**) Scheme of two "AT-less"

PKS modules on DepBC proteins that require a PPTase for carrier protein phosphopantetheinylation and a *trans* AT for polyketide chain elongation in FK228 biosynthesis. An inactive DH domain is drawn in light gray and labeled as DH^i. KS, AT, ACP, DH, KR and PPTase are standard abbreviations of domain/enzyme names that have been described in the text.

FIG 2. Reconstitution of FK228 biosynthesis in *E. coli* cells. (**a**) Scheme of a three-plasmid approach for reconstitution of FK228 biosynthetic pathway in *E. coli* BL21(DE3) cells. Only one cell containing one copy of each plasmid is drawn into picture for simplicity. A series of engineered strains were generated with different combinations of plasmids or different genes on plasmids (Table 1). (**b**) Examination of gene expression by semi-quantitative RT-PCR in engineered strain SW07 cultivated under aerobic conditions. 16S rDNA was amplified as an internal control.

FIG 3. Levels of FK228 produced by recombinant *E. coli* strains relative to the wild type strain of *Chromobacterium violaceum* no. 968 cultivated under aerobic conditions unless indicated by * for under anaerobic conditions. Data are mean values from duplicate experiments with error bars indicating standard deviation. Detailed strain information and FK228 levels are provided in Table 1 and Table S2.

1 TABLE 1. Bacterial strains and plasmids used in this study

Strains or plasmids	Description[a]	Source or reference
Chromobacterium violaceum		
no. 968 (=FERM BP-1968)	Wild type strain, FK228-producing, Ap^r $Thio^r$	IPOD[b]
Escherichia coli		
DH5α	General cloning host	Lab stock
S17-1	Host strain for interspecies conjugation	Lab stock
BL21(DE3)	Host strain for heterologous gene expression	Novagen
SW01	BL21(DE3) harboring Cosmid 18	This work
SW02	BL21(DE3) harboring Cosmid 18 and pBMTL-3-*depR*	This work
SW03	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*sfp*	This work
SW04	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*acpS*	This work
SW05	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*fabD1*	This work
SW06	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*fabD2*	This work
SW07	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*sfp-fabD1*	This work
SW08	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*sfp-fabD2*	This work
SW09	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*acpS-fabD1*	This work
SW10	BL21(DE3) harboring Cosmid 18, pBMTL-3-*depR* and pCDFDuet-1-*acpS-fabD2*	This work
Plasmids		
pGEM-3Zf	Ap^r, general cloning vector	Promega
pGEM-T Easy	Ap^r, general cloning vector	Promega
Cosmid 18	Ap^r, Kan^r, cosmid clone containing the FK228 biosynthetic gene cluster (*dep*) and flanking DNAs, shotgun sequenced	(9)
pBMTL-3	Cm^r, pBBR1 *ori*, broad host-range vector	(23)
pBMTL-3-*depR*	Cm^r, *depR* (with RBS from pET29a) cloned into pBMTL-3	(27)
pCDFDuet-1	Sm^r, CDF *ori*, dual expression vector	Novagen
pCDFDuet-1-*sfp*	*sfp* cloned into the MCS1 region of pCDFDuet-1	This study

pCDFDuet-1-*acpS*	*acpS* cloned into MCS1 region of pCDFDuet-1	This study
pCDFDuet-1-*fabD1*	*fabD1* cloned into MCS2 region of pCDFDuet-1	This study
pCDFDuet-1-*fabD2*	*fabD2* cloned into MCS2 region of pCDFDuet-1	This study
pCDFDuet-1-*sfp-fabD1*	*sfp* cloned into MCS1, *fabD1* cloned into MCS2 of pCDFDuet-1	This study
pCDFDuet-1-*sfp-fabD2*	*sfp* cloned into MCS1, *fabD2* cloned into MCS2 of pCDFDuet-1	This study
pCDFDuet-1-*acpS-fabD1*	*acpS* cloned into MCS1, *fabD1* cloned into MCS2 of pCDFDuet-1	This study
pCDFDuet-1-*acpS-fabD2*	*acpS* cloned into MCS1, *fabD2* cloned into MCS2 of pCDFDuet-1	This study

[a] Ap^r, ampicillin resistance; Cm^r, chloramphenicol resistance; Kan^r, kanamycin resistance; Sm^r, streptomycin resistance; $Thio^r$, thiostrepton resistance. MCS, multiple cloning site.

[b] IPOD, International Patent Organism Depositary, Tsukuba, Japan.

TABLE 2. Candidate genes identified through rapid genome sequencing and their involvement in FK228 biosynthesis

Gene Name[a]	Deduced Product and Predicted Function	Gene Dispensability	Involvement in FK228 Biosynthesis
AT-encoding genes			
fabD1	FabD1, likely involved in fatty acid biosynthesis	No	Positive
fabD2	FabD2, function unpredictable *a priori*	No	Positive
AT3	AT3, likely involved in cell-surface O-antigen biosynthesis	Yes	No
PPTase-encoding genes			
acpS	AcpS-type PPTase, likely involved in primary metabolite biosynthesis	No	No
sfp	Sfp-type PPTase, likely involved in secondary metabolite biosynthesis	No	Yes

[a] AT, acyltransferase; PPTase, phosphopantetheinyltransferase.

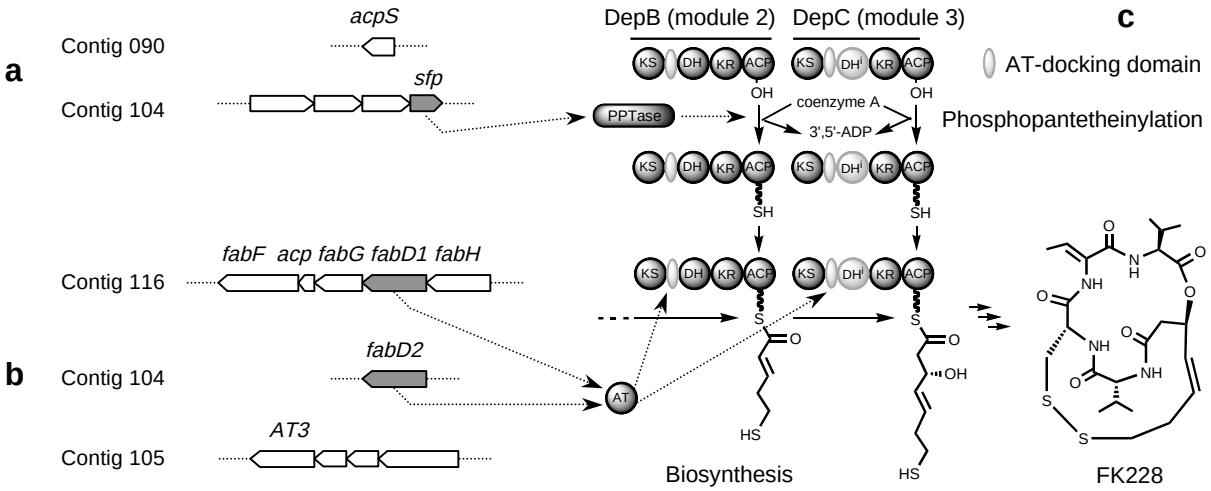


FIG 1. Identification of missing genes necessary for FK228 biosynthesis in *Chromobacterium violaceum* no. 968. (**a**) Local genetic map of two candidate PPTase-encoding genes, with the *sfp* gene product postulated to participate in phosphopantetheinylation of carrier proteins of the FK228 biosynthetic pathway. (**b**) Local genetic map of three candidate AT-encoding genes, with the *fabD1 and fabD2* gene products postulated to provide AT activities to complement the "AT-less" PKS modules on DepBC proteins for FK228 biosynthesis. (**c**) Scheme of two "AT-less" PKS modules on DepBC proteins that require a PPTase for carrier protein phosphopantetheinylation and a *trans* AT for polyketide chain elongation in FK228 biosynthesis. An inactive DH domain is drawn in light gray and labeled as DH^i. KS, AT, ACP, DH, KR and PPTase are standard abbreviations of domain/enzyme names that have been described in the text.

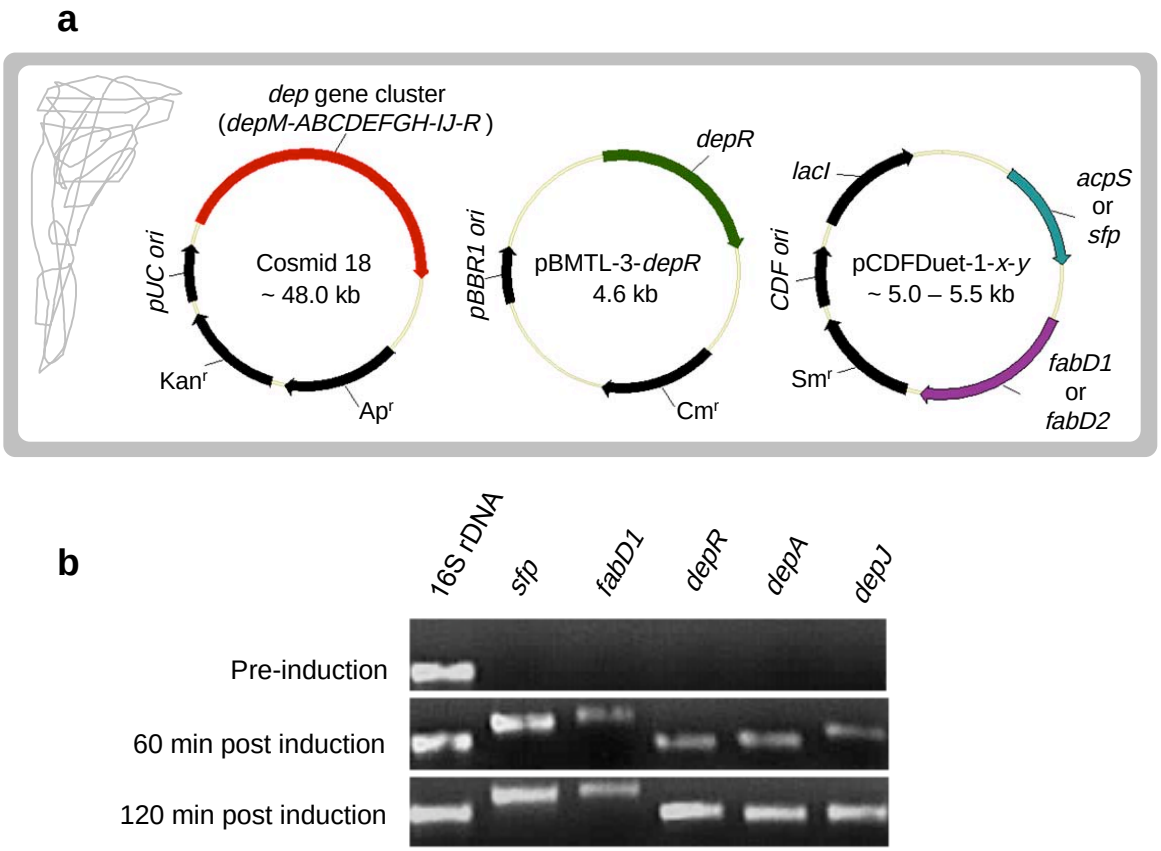


FIG 2. Reconstitution of FK228 biosynthesis in *E. coli* cells. (**a**) Scheme of a three-plasmid approach for reconstitution of FK228 biosynthetic pathway in *E. coli* BL21(DE3) cells. Only one cell containing one copy of each plasmid is drawn into picture for simplicity. A series of engineered strains were generated with different combinations of plasmids or different genes on plasmids (Table 1). (**b**) Examination of gene expression by semi-quantitative RT-PCR in engineered strain SW07 cultivated under aerobic conditions. 16S rDNA was amplified as an internal control.

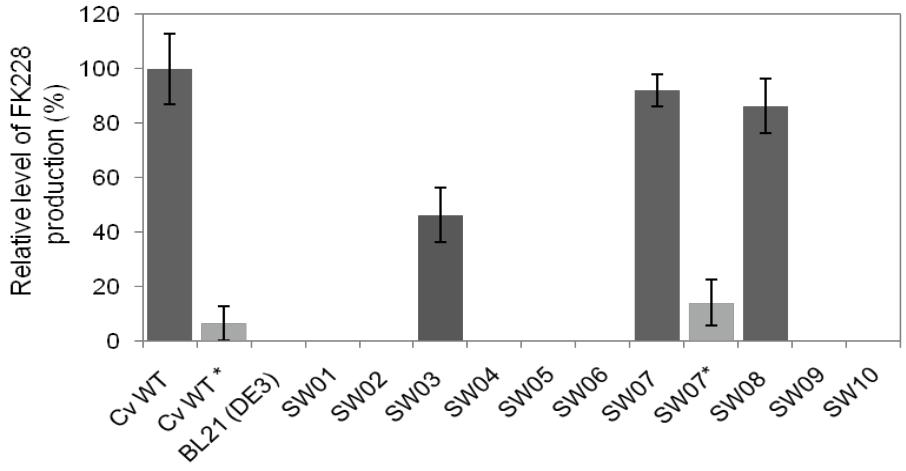


FIG 3. Levels of FK228 produced by recombinant *E. coli* strains relative to the wild type strain of *Chromobacterium violaceum* no. 968 cultivated under aerobic conditions unless indicated by * for under anaerobic conditions. Data are mean values from duplicate experiments with error bars indicating standard deviation. Detailed strain information and FK228 levels are provided in Table 1 and Table S2.